

# SCORE Search Results Details for Application 09961086 and Search Result 20080917\_142909\_us-09-961-086a-1.rup.

<a href="#">Score Home</a>	<a href="#">Retrieve Application</a>	<a href="#">SCORE System</a>	<a href="#">SCORE</a>	<a href="#">Comments /</a>
<a href="#">Page</a>	<a href="#">List</a>	<a href="#">Overview</a>	<a href="#">FAQ</a>	<a href="#">Suggestions</a>

This page gives you Search Results detail for the Application 09961086 and Search Result 20080917\_142909\_us-09-961-086a-1.rup.

[Go Back to previous page](#)

GenCore version 6.2.1  
Copyright (c) 1993 - 2008 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: September 18, 2008, 21:56:07 ; Search time 407 Seconds  
(without alignments)  
3112.639 Million cell updates/sec

Title: US-09-961-086A-1  
Perfect score: 3352  
Sequence: 1 MSSSNVEVFIPVSQGNTNGF.....MIVIFLTIAYLKLLFLKKYS 655

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5939836 seqs, 1934112985 residues

Total number of hits satisfying chosen parameters: 5939836

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_13.2:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	%		Query			Description
	Score	Match	Length	DB	ID	
1	3346	99.8	655	1	ABCG2_HUMAN	Q9unq0 homo sapien
2	3346	99.8	655	2	A8K1T5_HUMAN	A8k1t5 homo sapien
3	3225	96.2	655	2	A9UKW2_MACMU	A9ukw2 macaca mula
4	3223.5	96.2	654	1	ABCG2_MACMU	Q5mb13 macaca mula
5	3089	92.2	607	2	Q4W5I3_HUMAN	Q4w5i3 homo sapien
6	2890	86.2	658	2	Q09GP3_CAPHI	Q09gp3 capra hircu
7	2886	86.1	658	2	Q009B1_SHEEP	Q009b1 ovis aries
8	2870	85.6	658	2	A7E3T8_BOVIN	A7e3t8 bos taurus
9	2862	85.4	655	1	ABCG2_BOVIN	Q4gzt4 bos taurus
10	2859	85.3	658	2	Q32PJ1_BOVIN	Q32pj1 bos taurus
11	2849.5	85.0	656	1	ABCG2_PIG	Q8mib3 sus scrofa
12	2789	83.2	655	2	Q38JL0_CANFA	Q38jl0 canis famil
13	2762	82.4	657	1	ABCG2_MOUSE	Q7tms5 mus musculu
14	2754	82.2	657	1	ABCG2_RAT	Q80w57 rattus norv
15	2343	69.9	661	2	Q28BS4_XENTR	Q28bs4 xenopus tro
16	2288	68.3	661	2	A1L2M4_XENLA	A1l2m4 xenopus lae
17	2062	61.5	643	2	Q2Q447_DANRE	Q2q447 danio rerio
18	2042	60.9	655	2	A8IJF9_ONCMY	A8ijf9 oncorhynchu
19	1974.5	58.9	631	2	Q4SBP6_TETNG	Q4sbp6 tetraodon n
20	1787.5	53.3	650	2	Q8BKI5_MOUSE	Q8bki5 mus musculu
21	1786.5	53.3	650	1	ABCG3_MOUSE	Q99p81 mus musculu
22	1744.5	52.0	646	2	Q4KM08_RAT	Q4km08 rattus norv
23	1703.5	50.8	646	2	Q68HW7_RAT	Q68hw7 rattus norv
24	1663	49.6	613	2	Q2Q444_DANRE	Q2q444 danio rerio
25	1578.5	47.1	652	2	Q498U1_RAT	Q498u1 rattus norv
26	1473	43.9	634	2	Q08CU5_DANRE	Q08cu5 danio rerio
27	1469	43.8	634	2	Q2Q445_DANRE	Q2q445 danio rerio
28	1423	42.5	618	2	Q2Q446_DANRE	Q2q446 danio rerio
29	1422	42.4	618	2	A2BE75_DANRE	A2be75 danio rerio
30	1373	41.0	544	2	A7S071_NEMVE	A7s071 nematostell
31	1158	34.5	502	2	Q5U314_RAT	Q5u314 rattus norv
32	1038.5	31.0	457	2	Q4RBH3_TETNG	Q4rbh3 tetraodon n
33	1036.5	30.9	354	2	Q4SPA5_TETNG	Q4spa5 tetraodon n
34	940	28.0	1159	2	Q54T02_DICDI	Q54t02 dictyosteli
35	891.5	26.6	646	2	Q38AM7_9TRYP	Q38am7 trypanosoma
36	877	26.2	682	2	Q4DW41_TRYCR	Q4dw41 trypanosoma
37	875	26.1	619	2	A9VA57_MONBE	A9va57 monosiga br
38	872	26.0	645	2	A0CJS8_PARTE	A0cjs8 paramecium
39	870.5	26.0	607	2	Q22MH6_TETTH	Q22mh6 tetrahymena
40	866.5	25.9	827	2	A9UUE4_MONBE	A9uue4 monosiga br
41	864	25.8	1039	2	Q6BIH1_DEBHA	Q6bih1 debaryomyce
42	863.5	25.8	867	2	Q24CW4_TETTH	Q24cw4 tetrahymena
43	863	25.7	645	2	Q6BG61_PARTE	Q6bg61 paramecium
44	862.5	25.7	1006	2	A5DNC5_PICGU	A5dnc5 pichia guil
45	854.5	25.5	680	2	A4HPF5_LEIBR	A4hpf5 leishmania

## ALIGNMENTS

## RESULT 1

## ABCG2\_HUMAN

ID ABCG2\_HUMAN Reviewed; 655 AA.

AC Q9UNQ0; A0A1W3; O95374; Q53ZQ1; Q569L4; Q5YLG4; Q86V64; Q8IX16;

AC Q96LD6; Q96TA8; Q9BY73; Q9NUS0;

DT 24-JAN-2001, integrated into UniProtKB/Swiss-Prot.

DT 10-MAY-2005, sequence version 3.

DT 08-APR-2008, entry version 84.

DE ATP-binding cassette sub-family G member 2 (Placenta-specific ATP-binding cassette transporter) (Breast cancer resistance protein) (Mitoxantrone resistance-associated protein) (CD338 antigen) (CDw338).

GN Name=ABCG2; Synonyms=ABCP, BCRP, BCRP1, MXR;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

OC Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), VARIANTS GLU-166 AND SER-208, AND TISSUE SPECIFICITY.

RC TISSUE=Placenta;

RX MEDLINE=99065313; PubMed=9850061;

RA Allikmets R., Schriml L.M., Hutchinson A., Romano-Spica V., Dean M.;

RT "A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance.";

RL Cancer Res. 58:5337-5339(1998).

RN [2]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.

RC TISSUE=Mammary cancer;

RX MEDLINE=99080071; PubMed=9861027; DOI=10.1073/pnas.95.26.15665;

RA Doyle L.A., Yang W., Abruzzo L.V., Krogmann T., Gao Y., Rishi A.K., Ross D.D.;

RT "A multidrug resistance transporter from human MCF-7 breast cancer cells.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:15665-15670(1998).

RN [3]

RP ERRATUM.

RA Doyle L.A., Yang W., Abruzzo L.V., Krogmann T., Gao Y., Rishi A.K., Ross D.D.;

RA Ross D.D.;

RL Proc. Natl. Acad. Sci. U.S.A. 96:2569-2569(1999).

RN [4]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).

RA Kage K., Tsukahara S., Sugiyama T., Asada S., Ishikawa E., Tsuruo T., Sugimoto Y.;

RT "Breast cancer resistance protein constitutes a 140-kDa complex as a  
 RT homodimer.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
 RX MEDLINE=21201983; PubMed=11306452;  
 RA Komatani H., Kotani H., Hara Y., Nakagawa R., Matsumoto M.,  
 RA Arakawa H., Nishimura S.;  
 RT "Identification of breast cancer resistant protein/mitoxantrone  
 RT resistance/placenta-specific, ATP-binding cassette transporter as a  
 RT transporter of NB-506 and J-107088, topoisomerase I inhibitors with an  
 RT indolocarbazole structure.";  
 RL Cancer Res. 61:2827-2832(2001).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
 RX MEDLINE=21424790; PubMed=11533706; DOI=10.1038/nm0901-1028;  
 RA Zhou S., Schuetz J.D., Bunting K.D., Colapietro A.M., Sampath J.,  
 RA Morris J.J., Lagutina I., Grosveld G.C., Osawa M., Nakauchi H.,  
 RA Sorrentino B.P.;  
 RT "The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of  
 RT stem cells and is a molecular determinant of the side-population  
 RT phenotype.";  
 RL Nat. Med. 7:1028-1034(2001).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, AND VARIANTS GLU-166  
 RP AND SER-208.  
 RC TISSUE=Brain endothelium;  
 RX MEDLINE=22959505; PubMed=12958161; DOI=10.1096/fj.02-1131fje;  
 RA Zhang W., Mojsilovic-Petrovic J., Andrade M.F., Zhang H., Ball M.,  
 RA Stanimirovic D.B.;  
 RT "The expression and functional characterization of ABCG2 in brain  
 RT endothelial cells and vessels.";  
 RL FASEB J. 17:2085-2087(2003).  
 RN [8]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND VARIANT LYS-141.  
 RA Yoshikawa M., Yabuuchi H., Ikegami Y., Ishikawa T.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [9]  
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND VARIANT PRO-316.  
 RA Sudarikov A., Makarik T., Andreeff M.;  
 RT "Cell line K562 resistant to Hoechst 33342.";  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.  
 RN [10]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS MET-12; LYS-141;  
 RP HIS-296 AND THR-528.  
 RG SeattleSNPs program for genomic applications;  
 RL Submitted (SEP-2006) to the EMBL/GenBank/DDBJ databases.  
 RN [11]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2), AND VARIANT

RP LYS-141.  
 RC TISSUE=Pancreas, and PNS;  
 RX PubMed=15489334; DOI=10.1101/gr.2596504;  
 RG The MGC Project Team;  
 RT "The status, quality, and expansion of the NIH full-length cDNA  
 RT project: the Mammalian Gene Collection (MGC).";  
 RL Genome Res. 14:2121-2127(2004).  
 RN [12]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 198-655 (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 RN [13]  
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 294-655 (ISOFORM 1).  
 RX PubMed=9892175;  
 RA Miyake K., Mickley L., Litman T., Zhan Z., Robey R.W., Cristensen B.,  
 RA Brangi M., Greenberger L., Dean M., Fojo T., Bates S.E.;  
 RT "Molecular cloning of cDNAs which are highly overexpressed in  
 RT mitoxantrone-resistant cells: demonstration of homology to ABC

RT transport genes.";

RL Cancer Res. 59:8-13(1999).

RN [14]

RP REVIEW.

RX MEDLINE=21474438; PubMed=11590207;

RA Schmitz G., Langmann T., Heimerl S.;

RT "Role of ABCG1 and other ABCG family members in lipid metabolism.";

RL J. Lipid Res. 42:1513-1520(2001).

RN [15]

RP VARIANTS MET-12 AND LYS-141.

RX MEDLINE=22106379; PubMed=12111378; DOI=10.1007/s100380200041;

RA Iida A., Saito S., Sekine A., Mishima C., Kitamura Y., Kondo K.,

RA Harigae S., Osawa S., Nakamura Y.;

RT "Catalog of 605 single-nucleotide polymorphisms (SNPs) among 13 genes

RT encoding human ATP-binding cassette transporters: ABCA4, ABCA7, ABCA8,

RT ABCD1, ABCD3, ABCD4, ABCE1, ABCF1, ABCG1, ABCG2, ABCG4, ABCG5, and

RT ABCG8.";

RL J. Hum. Genet. 47:285-310(2002).

RN [16]

RP VARIANTS LEU-431 AND LEU-489.

RX PubMed=15618737; DOI=10.2133/dmpk.18.212;

RA Itoda M., Saito Y., Shirao K., Minami H., Ohtsu A., Yoshida T.,

RA Saijo N., Suzuki H., Sugiyama Y., Ozawa S., Sawada J.;

RT "Eight novel single nucleotide polymorphisms in ABCG2/BCRP in Japanese

RT cancer patients administered irinotacan.";

RL Drug Metab. Pharmacokinet. 18:212-217(2003).

RN [17]

RP VARIANTS MET-12; LYS-141; LEU-206 AND TYR-590.

RX PubMed=12544509; DOI=10.1097/00008571-200301000-00004;

RA Zamber C.P., Lamba J.K., Yasuda K., Farnum J., Thummel K.,

RA Schuetz J.D., Schuetz E.G.;

RT "Natural allelic variants of breast cancer resistance protein (BCRP)

RT and their relationship to BCRP expression in human intestine.";

RL Pharmacogenetics 13:19-28(2003).

RN [18]

RP EFFECT OF THE VARIANTS MET-12; LYS-141 AND ASN-620 ON TRANSPORT.

RX PubMed=15838659; DOI=10.1007/s00280-004-0931-x;

RA Morisaki K., Robey R.W., Oezvegy-Laczka C., Honjo Y., Polgar O.,

RA Steadman K., Sarkadi B., Bates S.E.;

RT "Single nucleotide polymorphisms modify the transporter activity of

RT ABCG2.";

RL Cancer Chemother. Pharmacol. 56:161-172(2005).

RN [19]

RP SUBCELLULAR LOCATION, GLYCOSYLATION AT ASN-596, AND MUTAGENESIS OF

RP ASN-418; ASN-557 AND ASN-596.

RX PubMed=15807535; DOI=10.1021/bi0479858;

RA Diop N.K., Hrycyna C.A.;

RT "N-linked glycosylation of the human ABC transporter ABCG2 on

RT asparagine 596 is not essential for expression, transport activity, or

RT trafficking to the plasma membrane.";  
RL Biochemistry 44:5420-5429(2005).  
RN [20]  
RP MUTAGENESIS OF LYS-86, SUBCELLULAR LOCATION, AND HOMODIMERIZATION.  
RX PubMed=15769853; DOI=10.1242/jcs.01729;  
RA Henriksen U., Gether U., Litman T.;  
RT "Effect of Walker A mutation (K86M) on oligomerization and surface  
RT targeting of the multidrug resistance transporter ABCG2.";  
RL J. Cell Sci. 118:1417-1426(2005).  
RN [21]  
RP MUTAGENESIS OF ARG-482.

Query Match 99.8%; Score 3346; DB 1; Length 655;  
Best Local Similarity 99.8%; Pred. No. 5e-211;  
Matches 654; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Db	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
Db	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Db	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
Db	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
Qy	241	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Db	241	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Qy	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Db	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Qy	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Db	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Qy	421	TGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Db	421	TGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Qy	481	MTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540

```

      | |||||
Db      481 MRMLPSIIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL 540
Qy      541 MTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN 600
      | |||||
Db      541 MTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN 600
Qy      601 NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655
      | |||||
Db      601 NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655
```

RESULT 2

A8K1T5\_HUMAN

```
ID      A8K1T5_HUMAN              Unreviewed;          655 AA.
AC      A8K1T5;
DT      04-DEC-2007, integrated into UniProtKB/TrEMBL.
DT      04-DEC-2007, sequence version 1.
DT      08-APR-2008, entry version 5.
DE      cDNA FLJ76761, highly similar to Homo sapiens ATP-binding cassette,
DE      sub-family G (WHITE), member 2(ABCG2), mRNA.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC      Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Hippocampus;
RA      Wakamatsu A., Yamamoto J., Kimura K., Ishii S., Watanabe K.,
RA      Sugiyama A., Murakawa K., Kaida T., Tsuchiya K., Fukuzumi Y.,
RA      Kumagai A., Oishi Y., Yamamoto S., Ono Y., Komori Y., Yamazaki M.,
RA      Kisu Y., Nishikawa T., Sugano S., Nomura N., Isogai T.;
RT      "NEDO human cDNA sequencing project.";
RL      Submitted (OCT-2007) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: Belongs to the ABC transporter family.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AK290000; BAF82689.1; -; mRNA.
DR      RefSeq; NP_004818.2; -.
DR      UniGene; Hs.480218; -.
DR      GeneID; 9429; -.
DR      GO; GO:0005524; F:ATP binding; IEA:InterPro.
DR      GO; GO:0016887; F:ATPase activity; IEA:InterPro.
DR      InterPro; IPR003593; AAA+_ATPase_core.
DR      InterPro; IPR013525; ABC_2_trans.
DR      InterPro; IPR003439; ABC_transp_like.
```



DR Pfam; PF01061; ABC2\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
PE 2: Evidence at transcript level;  
KW ATP-binding; Membrane; Nucleotide-binding; Transmembrane; Transport.  
SQ SEQUENCE 655 AA; 72314 MW; A8AF66B96034C5A8 CRC64;

Query Match 99.8%; Score 3346; DB 2; Length 655;  
Best Local Similarity 99.8%; Pred. No. 5e-211;  
Matches 654; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Db	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
Db	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Db	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
Db	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
Qy	241	SIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Db	241	SIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Qy	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Db	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Qy	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Db	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Qy	421	TGIQNRAGVLFFLTNTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Db	421	TGIQNRAGVLFFLTNTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Qy	481	MTMLPSIIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540
Db	481	MRMLPSIIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540

Qy 541 MTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN 600  
|||||  
Db 541 MTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN 600

Qy 601 NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655  
|||||  
Db 601 NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655

RESULT 3

A9UKW2\_MACMU

ID A9UKW2\_MACMU Unreviewed; 655 AA.

AC A9UKW2;

DT 05-FEB-2008, integrated into UniProtKB/TrEMBL.

DT 05-FEB-2008, sequence version 1.

DT 08-APR-2008, entry version 2.

DE ATP-binding cassette transporter sub-family G member 2.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

OC Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.

OX NCBI\_TaxID=9544;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Nakanishi T., Tsang A., Cheng X., Ross D.D., MacVittie T., Takebe N.;

RT "cDNA cloning and functional analysis of rhesus monkey ATP-binding

RT cassette transporter, BCRP/ABCG2.";

RL Submitted (DEC-2004) to the EMBL/GenBank/DDBJ databases.

CC -----

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NoDerivs License

CC -----

DR EMBL; AY864772; AAX56948.1; -; mRNA.

DR InterPro; IPR003593; AAA+\_ATPase\_core.

DR InterPro; IPR013525; ABC\_2\_trans.

DR InterPro; IPR003439; ABC\_transp\_like.

DR Pfam; PF01061; ABC2\_membrane; 1.

DR Pfam; PF00005; ABC\_tran; 1.

DR ProDom; PD000006; ABC\_transporter; 1.

DR SMART; SM00382; AAA; 1.

DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.

PE 2: Evidence at transcript level;

KW ATP-binding.

SQ SEQUENCE 655 AA; 72601 MW; CE1DEABF5C0648DB CRC64;

Query Match 96.2%; Score 3225; DB 2; Length 655;

Best Local Similarity 96.2%; Pred. No. 4.6e-203;

Matches 630; Conservative 7; Mismatches 18; Indels 0; Gaps 0;



DT 21-JUN-2005, integrated into UniProtKB/Swiss-Prot.  
 DT 01-FEB-2005, sequence version 1.  
 DT 15-JAN-2008, entry version 25.  
 DE ATP-binding cassette sub-family G member 2 (CD338 antigen).  
 GN Name=ABCG2;  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 OC Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.  
 RC TISSUE=Kidney;  
 RX PubMed=15516692; DOI=10.1074/jbc.M409796200;  
 RA Ueda T., Brenner S., Malech H.L., Langemeijer S.M., Perl S., Kirby M.,  
 RA Phang O.A., Krouse A.E., Donahue R.E., Kang E.M., Tisdale J.F.;  
 RT "Cloning and functional analysis of the rhesus macaque ABCG2 gene.  
 RT Forced expression confers an SP phenotype among hematopoietic stem  
 RT cell progeny in vivo."  
 RL J. Biol. Chem. 280:991-998(2005).  
 CC -!- FUNCTION: Xenobiotic transporter that may play an important role  
 CC in the exclusion of xenobiotics from the brain. May be involved in  
 CC brain-to-blood efflux (By similarity). When overexpressed, the  
 CC transfected cells become resistant to mitoxantrone. Overexpression  
 CC in bone marrow stem cells does not interfere with hematopoietic  
 CC stem cell maturation and increases the number of SP cells.  
 CC -!- SUBUNIT: Monomer or homodimer; disulfide-linked (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cell membrane; Multi-pass membrane protein  
 CC (By similarity).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCG (White)  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 ABC transporter domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AY841878; AAW28901.1; -; mRNA.  
 DR RefSeq; NP\_001028091.1; -.  
 DR UniGene; Mmu.3144; -.  
 DR Ensembl; ENSMMUG00000008797; Macaca mulatta.  
 DR GeneID; 574307; -.  
 DR GO; GO:0005886; C:plasma membrane; IEA:UniProtKB-SubCell.  
 DR InterPro; IPR003593; AAA+\_ATPase\_core.  
 DR InterPro; IPR013525; ABC\_2\_trans.  
 DR InterPro; IPR003439; ABC\_transp\_like.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.

Query Match 96.2%; Score 3223.5; DB 1; Length 654;  
Best Local Similarity 96.5%; Pred. No. 5.7e-203;  
Matches 632; Conservative 7; Mismatches 15; Indels 1; Gaps 1;

[http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7\\_142909\\_us-09-961-086a-1.rup&ItemType=4&startByte=0](http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7_142909_us-09-961-086a-1.rup&ItemType=4&startByte=0) (13 of 40)9/22/2008 12:01:40 PM

Qy	241	SIHQPRYSIFKLFDSLTLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Db	241	SIHQPRYSIFKLFDSLTLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Qy	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
		:    :     :	
Db	301	DSTAVALNREEDFKATEIIEPSKRDKPLVEKLAEIYVDSSFYKETKAELHQLSGGE-KKK	359
Qy	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
		:	
Db	360	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVILGLVIGAIYFGLNDS	419
Qy	421	TGIQNRAGVLFFLTNTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Db	420	TGIQNRAGVLFFLTNTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFFGKLLSDLLP	479
Qy	481	MTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540
		:	
Db	480	MRMLPSIIFTCIVYFMLGLKPTADAFFIMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	539
Qy	541	MTICFVFMMIFSGLLVNLTITIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN	600
Db	540	MTICFVFMMIFSGLLVNLTITIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATVN	599
Qy	601	NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
Db	600	NTCNYATCTGEEYLAKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	654

RESULT 5

Q4W5I3\_HUMAN

ID	Q4W5I3_HUMAN	Unreviewed;	607 AA.
AC	Q4W5I3;		
DT	05-JUL-2005,	integrated into UniProtKB/TrEMBL.	
DT	05-JUL-2005,	sequence version 1.	
DT	08-APR-2008,	entry version 21.	
DE	Putative uncharacterized protein ABCG2 (Fragment).		
GN	Name=ABCG2;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;		
OC	Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Spalding L., Kozlowicz A., Abbott S.;		
RT	"The sequence of Homo sapiens BAC clone RP11-147K6.";		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.		
RN	[2]		

RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.H.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RA Wilson R.K.;  
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AC097484; AAY40902.1; -; Genomic\_DNA.  
DR UniGene; Hs.480218; -.  
DR Ensembl; ENSG00000118777; Homo sapiens.  
DR HGNC; HGNC:74; ABCG2.  
DR ArrayExpress; Q4W5I3; -.  
DR GO; GO:0016021; C:integral to membrane; IEA:UniProtKB-KW.  
DR GO; GO:0005524; F:ATP binding; IEA:InterPro.  
DR GO; GO:0016887; F:ATPase activity; IEA:InterPro.  
DR GO; GO:0006810; P:transport; IEA:UniProtKB-KW.  
DR InterPro; IPR003593; AAA+\_ATPase\_core.  
DR InterPro; IPR013525; ABC\_2\_trans.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR Pfam; PF01061; ABC2\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
PE 2: Evidence at transcript level;  
KW ATP-binding; Membrane; Nucleotide-binding; Transmembrane; Transport.  
FT NON\_TER 607 607  
SQ SEQUENCE 607 AA; 66800 MW; 27124123FAD451DC CRC64;

Query Match 92.2%; Score 3089; DB 2; Length 607;  
Best Local Similarity 99.8%; Pred. No. 3.6e-194;  
Matches 606; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Db	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
Db	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120

Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Db	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
Db	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
Qy	241	SIHQPRYSIFKLFDSTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Db	241	SIHQPRYSIFKLFDSTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Qy	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Db	301	DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Qy	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Db	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Qy	421	TGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Db	421	TGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Qy	481	MTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540
Db	481	MRMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540
Qy	541	MTICFVFMMIFSGLLVNLTITIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN	600
Db	541	MTICFVFMMIFSGLLVNLTITIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN	600
Qy	601	NPCNYAT	607
Db	601	NPCNYAT	607

RESULT 6  
Q09GP3\_CAPHI  
ID Q09GP3\_CAPHI Unreviewed; 658 AA.  
AC Q09GP3;  
DT 17-OCT-2006, integrated into UniProtKB/TrEMBL.  
DT 17-OCT-2006, sequence version 1.  
DT 08-APR-2008, entry version 12.  
DE ATP-binding cassette sub-family G member 2.  
GN Name=ABCG2;  
OS Capra hircus (Goat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
OC Pecora; Bovidae; Caprinae; Capra.  
OX NCBI\_TaxID=9925;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Wu H., Luo J., Zhang L.;  
RT "Cloning and sequence analyses of ABCG2 gene differentially expressed  
RT in mammary gland at two lactation stages of Xinong Saanen goat.";  
RL Submitted (AUG-2006) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; DQ904356; ABI73985.1; -; mRNA.  
DR GO; GO:0016021; C:integral to membrane; IEA:UniProtKB-KW.  
DR GO; GO:0005524; F:ATP binding; IEA:InterPro.  
DR GO; GO:0016887; F:ATPase activity; IEA:InterPro.  
DR GO; GO:0006810; P:transport; IEA:UniProtKB-KW.  
DR InterPro; IPR003593; AAA+\_ATPase\_core.  
DR InterPro; IPR013525; ABC\_2\_trans.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR Pfam; PF01061; ABC2\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
PE 2: Evidence at transcript level;  
KW ATP-binding; Membrane; Nucleotide-binding; Transmembrane; Transport.  
SQ SEQUENCE 658 AA; 73200 MW; C8BD65DF4E877D62 CRC64;

Query Match 86.2%; Score 2890; DB 2; Length 658;  
Best Local Similarity 85.2%; Pred. No. 5e-181;  
Matches 559; Conservative 43; Mismatches 52; Indels 2; Gaps 2;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTGAVLSFHNICYRVKLKSGFLPCRKPVE	60
		::      :: :            :      ::   :: :      :	
Db	4	MSSNSYEVCIPMSK-KPNGIPETTSKDLQTLTEGAVLSFHDICYRVKVKTGFLLCRKTIE	62
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:    :     :     :     :     :     :	
Db	63	KEILANINGVMKPGLNAILGPTGGGKSSLLDILAARKDPHGLSGDVLINGAPRPANFKCN	122
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATMTNHEKNERINRVIQELGLDKVADSKVGT	180
		:     :     :     :     :     :     :	
Db	123	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTNYEKNERINKVIQELGLDKVADSKVGT	182
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
		:     :     :     :     :     :     :	

Db	183	QFIRGVSGGERKRTSIAMELITDPSILFLDEPTTGLDSSTANAVLLLLLKRM SKQGRTIIF	242
Qy	241	SIHQPRYSIFKLFDSLTL LASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
Db	243	SIHQPRYSIFKLFDSLTL LASGRLMFHGPAQEALGYFEDIGFHCEPYNNPADFFLDIING	302
Qy	301	DSTAVALNREE-DFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:        :   :                     :         :     : :	
Db	303	DSSAVVLNREDSDEAKETEEPSKNDTSLIEKLAEFYVNSSFYKETKVELDKFSGEQRRK	362
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKND	419
		: : :       :                         :               : :	
Db	363	KLSSYKEITYATSFCHQLKWISKRSFKNLLGNPQASIAQLIVTVFLGLVIGAIIFYDLKND	422
Qy	420	STGIQNRAGVLFFLT TNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL	479
		:	
Db	423	PSGIQNRAGVLFFLT TNQCFSSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSDLL	482
Qy	480	PMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL	539
		:         :     :                             :	
Db	483	PMRMLPSIIFTCITYFLGLKPKVEAFFIMMFTLMMVAYSASSMALAIAAGQSVVSIATL	542
Qy	540	LMTICFVFM MIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATG	599
		:                   :	
Db	543	LMTISFVFM MIFSGLLVNLKTIGAWLSWLQYLSIPRYGYAALQHNEFLGQNFCPGLNVT A	602
Qy	600	NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
		:           :       :                                   :	
Db	603	NNTCSYAICTGEEFLTNQGIDISPWGLWKNHVALACMIVIFLTIAYLKLLFLKKFS	658

RESULT 7  
Q009B1\_SHEEP  
ID Q009B1\_SHEEP Unreviewed; 658 AA.  
AC Q009B1;  
DT 14-NOV-2006, integrated into UniProtKB/TrEMBL.  
DT 14-NOV-2006, sequence version 1.  
DT 08-APR-2008, entry version 13.  
DE ATP-binding cassette sub-family G member 2.  
GN Name=ABCG2;  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
OC Pecora; Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Duncan E.J., Dodds K.G., Henry H.M., Thompson M.P., Phua S.H.;  
RT "Cloning, mapping and association studies of the ovine ABCG2 gene with

Query Match 86.1%; Score 2886; DB 2; Length 658;  
Best Local Similarity 85.2%; Pred. No. 9.2e-181;  
Matches 559; Conservative 41; Mismatches 54; Indels 2; Gaps 2;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKCLKSGFLPCRKPVE	60
		::      ::              :                 :       :	
Db	4	MSSNSYEVSIPMSK-KPNGIPETTSKDLQTLTEGAVLSFHNICYRVKVKTGFLLCRKTI	62
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:    :                          :	
Db	63	KEILANINGVMKPGLNAILGPTGGGKSSLLDILAARKDPHGLSGDVLINGAPRPANFKCN	122
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
		:           :	
Db	123	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTNIEKNERINKVIQELGLDKVADSKVGT	182
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
Db	183	QFIRGVSGGERKRTSIAMEELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	242
Qy	241	SIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:	

Db	243	SIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFEDIGFHCEPYNNPADFFLDIING	302
Qy	301	DSTAVALNREE-DFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:        :   :                     :       :	
Db	303	DSSAVVLNREDSDEAKETEEPSKNDTSLIEKLAGFYVNSSFYKETKVELDKFSGERRRK	362
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKND	419
		: : :       :             :                 :	
Db	363	KLSSYKEITYATSFCHQLKWISKRSFKNLLGNPQASIAQLIVTVFLGLVIGAIFYDLKND	422
Qy	420	STGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL	479
		:	
Db	423	PSGIQNRAGVLFFLTNNQCFSSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSDLL	482
Qy	480	PMTMLPSIIFTTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL	539
		:         :       :                     :	
Db	483	PMRMLPSIIFTTCITYFLLGLKPKVEAFFIMMFTLMMVAYSASSMALAIAAGQSVVSIATL	542
Qy	540	LMTICFVFMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATG	599
		:                 :	
Db	543	LMTISFVFMIFSGLLVNLKTIGAWLSWLQYLSIPRYGYAALQHNEFLGQNFCPGLNVTA	602
Qy	600	NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
		:           :       :                             :	
Db	603	NNTCSYAICTGEEFLTNQGIDISPWGLWKNHVALACMIVIFLTIAYLKLLFLKKFS	658

## RESULT 8

## A7E3T8\_BOVIN

ID	A7E3T8_BOVIN	Unreviewed;	658 AA.
AC	A7E3T8;		
DT	11-SEP-2007, integrated into UniProtKB/TrEMBL.		
DT	11-SEP-2007, sequence version 1.		
DT	08-APR-2008, entry version 7.		
DE	ATP-binding cassette, sub-family G, member 2.		
GN	Name=ABCG2;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;		
OC	Pecora; Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Pooled;		
RX	MEDLINE=22135956; PubMed=12140684; DOI=10.1007/s00335-001-2145-4;		
RA	Sonstegard T.S., Capuco A.V., White J., Van Tassell C.P., Connor E.E.,		
RA	Cho J., Sultana R., Shade L., Wray J.E., Wells K.D., Quackenbush J.;		
RT	"Analysis of bovine mammary gland EST and functional annotation of the		
RT	Bos taurus gene index.";		

Query Match 85.6%; Score 2870; DB 2; Length 658;  
Best Local Similarity 84.5%; Pred. No. 1e-179;  
Matches 554; Conservative 45; Mismatches 55; Indels 2; Gaps 2;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKCLKSGFLPCRKPVE	60
		::      ::            :            :: :      :	
Db	4	MSSNSYEVSIPMSK-KLNGIPETTSKDLQTLTEGAVLSFHNICYRVKVKTGFLLCRKTIE	62
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPGLSGDVLINGAPRPANFKCN	120
		:    :                 :	
Db	63	KEILANINGVMKPGLNAILGPTGGGKSSLLDILAARKDPHGLSGDVLINGAPRPANFKCN	122

Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
		: :     :	
Db	123	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTSYEKNERINKVIQELGLDKVADSKVGT	182
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
		: :     :	
Db	183	QFIRGVSGGERKRTSIAMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	242
Qy	241	SIHQPRYSIFKLFDSTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		: :     :	
Db	243	SIHQPRYSIFKLFDSTLLASGRLMFHGPAQEALGYFGAIGFHCEPYNNPADFFLDIING	302
Qy	301	DSTAVALNREE-DFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:        :                     :           :     :	
Db	303	DSSAVVLNREDIGDEANETEEPSKKDTPLIEKLAEFYVNSSFFKETKVELDKFSGDQRRK	362
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKND	419
		:     :           :     :           :           :	
Db	363	KLPSYKEVTYATSFCHQLKWISRRSFKNLLGNPQASIAQLIVTVFLGLVIGAIIFYDLKND	422
Qy	420	STGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL	479
		: :     :	
Db	423	PAGIQNRAGVLFFLTNQCFSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSDLL	482
Qy	480	PMTMLPSIIFTTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL	539
		:         :     :                     :	
Db	483	PMRMLPSIIFTTCITYFLLGLKPKVEAFFIMMLTLMMVAYSASSMALAIAAGQSVVSIATL	542
Qy	540	LMTICFVFMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATG	599
		:                 :	
Db	543	LMTISFVFMIFSGLLVNLKTVVPWLSWLQYLSIPRYGYAALQHNEFLGQNFCPGLNVTT	602
Qy	600	NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
		:           :         :                           :	
Db	603	NNTCSYAICTGEEFLTNOGIDISPWGLWKNHVALACMIVIFLTIAYLKLLFLKKFS	658

RESULT 9

ABCG2\_BOVIN

ID	ABCG2_BOVIN	Reviewed;	655 AA.
AC	Q4GZT4;		
DT	27-JUN-2006, integrated into UniProtKB/Swiss-Prot.		
DT	27-JUN-2006, sequence version 2.		
DT	15-JAN-2008, entry version 24.		
DE	ATP-binding cassette sub-family G member 2 (CD338 antigen).		
GN	Name=ABCG2;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 OC Pecora; Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANT SER-578.  
 RC STRAIN=Holstein;  
 RX PubMed=15998908; DOI=10.1101/gr.3806705;  
 RA Cohen-Zinder M., Seroussi E., Larkin D.M., Looor J.J.,  
 RA Everts-van der Wind A., Lee J.-H., Drackley J.K., Band M.R.,  
 RA Hernandez A.G., Shani M., Lewin H.A., Weller J.I., Ron M.;  
 RT "Identification of a missense mutation in the bovine ABCG2 gene with a  
 RT major effect on the QTL on chromosome 6 affecting milk yield and  
 RT composition in Holstein cattle."  
 RL Genome Res. 15:936-944(2005).  
 CC -!- FUNCTION: Xenobiotic transporter that may play an important role  
 CC in the exclusion of xenobiotics from the brain. May be involved in  
 CC brain-to-blood efflux (By similarity).  
 CC -!- SUBUNIT: Monomer or homodimer; disulfide-linked (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cell membrane; Multi-pass membrane protein.  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCG (White)  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 ABC transporter domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AJ871176; CAI38796.1; ALT\_INIT; Genomic\_DNA.  
 DR UniGene; Bt.51973; -.  
 DR Ensembl; ENSBTAG00000017704; Bos taurus.  
 DR GO; GO:0005886; C:plasma membrane; IEA:UniProtKB-SubCell.  
 DR InterPro; IPR003593; AAA+ATPase\_core.  
 DR InterPro; IPR013525; ABC\_2\_trans.  
 DR InterPro; IPR003439; ABC\_transp\_like.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 PE 3: Inferred from homology;  
 KW ATP-binding; Glycoprotein; Membrane; Nucleotide-binding; Polymorphism;  
 KW Transmembrane; Transport.  
 FT CHAIN 1 655 ATP-binding cassette sub-family G member  
 FT 2.  
 FT /FTId=PRO\_0000244032.  
 FT TOPO\_DOM 1 395 Cytoplasmic (Potential).  
 FT TRANSMEM 396 416 Potential.  
 FT TOPO\_DOM 417 428 Extracellular (Potential).  
 FT TRANSMEM 429 449 Potential.

FT	TOPO_DOM	450	477	Cytoplasmic (Potential).
FT	TRANSMEM	478	498	Potential.
FT	TOPO_DOM	499	506	Extracellular (Potential).
FT	TRANSMEM	507	527	Potential.
FT	TOPO_DOM	528	535	Cytoplasmic (Potential).
FT	TRANSMEM	536	556	Potential.
FT	TOPO_DOM	557	630	Extracellular (Potential).
FT	TRANSMEM	631	651	Potential.
FT	TOPO_DOM	652	655	Cytoplasmic (Potential).
FT	DOMAIN	36	285	ABC transporter.
FT	NP_BIND	79	86	ATP (Potential).
FT	CARBOHYD	596	596	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	600	600	N-linked (GlcNAc. . .) (Potential).
FT	VARIANT	578	578	Y -> S (polymorphism affecting milk fat
FT				and protein concentration).
SQ	SEQUENCE	655 AA;	72725 MW;	8F1AD75742AD236E CRC64;

Query Match 85.4%; Score 2862; DB 1; Length 655;  
Best Local Similarity 84.3%; Pred. No. 3.5e-179;  
Matches 553; Conservative 45; Mismatches 56; Indels 2; Gaps 2;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKCLKSGFLPCRKPVE	60
		::      : :            :      : :       :	
Db	1	MSSNSYEVSIPMSK-KLNGIPETTSKDLQTLTEGAVLSFHNICYRVKVKTGFLLCRKTIE	59
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:    :     :     :     :     :     :	
Db	60	KEILANINGVMKPGLNAILGPTGGGKSSLLDILAARKDPHGLSGDVLINGAPRPANFKCN	119
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
		:     :     :     :     :     :     :	
Db	120	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTSYEKNERINKVIQELGLDKVADSKVGT	179
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
		:     :     :     :     :     :     :	
Db	180	QFIRGVSGGERKRTSIAMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	239
Qy	241	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:     :     :     :     :     :     :	
Db	240	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFGAIGRCEPYNNPADFFLDIING	299
Qy	301	DSTAVALNREE-DFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:       : :        :       :        :    :::	
Db	300	DSSAVVLNREDIGDEANETEEPSKKDTPLIEKLAEFYVNSSFYKETKVELDKFSGDQRRK	359
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKND	419
		: :  ::       : : :     :          : :	
Db	360	KLPSYKEVTYATSFCHQLKWISRRSFKNLLGNPQASIAQLIVTVFLGLVIGAIIFYDLKND	419



```
Qy      420 STGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSTYFLGKLLSDLL 479
          |||
Db      420 PAGIQNRAGVLFFLTNNQCFSSVSAVELLVVEKKLFIHEYISGYRVSSTYFFGKLLSDLL 479

Qy      480 PMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL 539
          || |||
Db      480 PMRMLPSIIFTCITYFLLGLKPKVEAFFIMMLTLMMVAYSASSMALAIAAGQSVVSIATL 539

Qy      540 LMTICFVFMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFPCPLNATG 599
          ||| |||
Db      540 LMTISFVFMIFSGLLVNLKTVVPWLSWLQYLSIPRYGYAALQHNEFLGQNFPCPLNVT 599

Qy      600 NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655
          || |
Db      600 NNTCSYAICTGEEFLTNQGIDISPWGLWKNHVALACMIVIFLTIAYLKLLFLKKFS 655
```

## RESULT 10

## Q32PJ1\_BOVIN

```
ID      Q32PJ1_BOVIN                      Unreviewed;          658 AA.
AC      Q32PJ1;
DT      06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT      19-SEP-2006, sequence version 2.
DT      08-APR-2008, entry version 36.
DE      ATP-binding cassette, sub-family G (WHITE), member 2.
GN      Name=ABCG2;
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC      Pecora; Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=Crossbred x Angus; TISSUE=Ileum;
RA      Moore S., Alexander L., Brownstein M., Guan L., Lobo S., Meng Y.,
RA      Tanaguchi M., Wang Z., Yu J., Prange C., Schreiber K., Shenmen C.,
RA      Wagner L., Bala M., Barbazuk S., Barber S., Babakaiff R., Beland J.,
RA      Chun E., Del Rio L., Gibson S., Hanson R., Kirkpatrick R., Liu J.,
RA      Matsuo C., Mayo M., Santos R.R., Stott J., Tsai M., Wong D.,
RA      Siddiqui A., Holt R., Jones S.J., Marra M.A.;
RL      Submitted (OCT-2005) to the EMBL/GenBank/DDBJ databases.
CC      -!- SIMILARITY: Belongs to the ABC transporter family.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; BC108097; AAI08098.2; -; mRNA.
DR      RefSeq; NP_001032555.2; -.
DR      UniGene; Bt.51973; -.
```

DR Ensembl; ENSBTAG00000017704; Bos taurus.  
 DR GeneID; 536203; -.  
 DR KEGG; bta:536203; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA:UniProtKB-KW.  
 DR GO; GO:0005524; F:ATP binding; IEA:InterPro.  
 DR GO; GO:0016887; F:ATPase activity; IEA:InterPro.  
 DR GO; GO:0006810; P:transport; IEA:UniProtKB-KW.  
 DR InterPro; IPR003593; AAA+\_ATPase\_core.  
 DR InterPro; IPR013525; ABC\_2\_trans.  
 DR InterPro; IPR003439; ABC\_transp\_like.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 PE 2: Evidence at transcript level;  
 KW ATP-binding; Membrane; Nucleotide-binding; Transmembrane; Transport.  
 SQ SEQUENCE 658 AA; 73113 MW; 53DB7AAF29B6202A CRC64;

Query Match 85.3%; Score 2859; DB 2; Length 658;  
 Best Local Similarity 84.1%; Pred. No. 5.5e-179;  
 Matches 552; Conservative 46; Mismatches 56; Indels 2; Gaps 2;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
		::      ::            :      : :      :	
Db	4	MSSNSYEVSIPI MSK-KLNGIPETTSKDLQTLTEGAVLSFHNICYRVKVKTGFLLCRKTIE	62
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:    :     :     :     :     :     :	
Db	63	KEILANINGVMKPGLNAILGPTGGGKSSLLDILAARKDPHGLSGDVLINGAPRPANFKCN	122
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
		:     :     :     :     :     :     :	
Db	123	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTMTSYEKNERINKVIQELGLDKVADSKVGT	182
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
		:     :     :     :     :     :     :	
Db	183	QFIRGVSGGERKRTSIAMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	242
Qy	241	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:     :     :     :     :     :     :	
Db	243	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFGAIGFRCEPYNNPADFFLDIING	302
Qy	301	DSTAVALNREE-DFKATEIIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:       : :       :       :         :     ::	
Db	303	DSSAVVLNREDIGDEANETEEPSKKDTPLIEKLAEFYVNSSFYKETKVELDKFSGDQRRK	362
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKND	419
		: :  ::       : : :     :    :            : :	

Db 363 KLPSYKEVTYATSFCHQLKWISRRSFKNLLGNPQSSIAQLIVTVFLGLVIGAIFYDLKND 422

Qy 420 STGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL 479  
|||||

Db 423 PAGIQNRAGVLFFLTNNQCFSSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSDLL 482

Qy 480 PMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL 539  
|| |||:|||||:||||| |||:|||||

Db 483 PMRMLPSIIFTCITYFLLGLKPKVEAFFIMMLTLMMVAYSASSMALAIAAGQSVVSIATL 542

Qy 540 LMTICFVFMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATG 599  
|||| |||: ||| |||: ||| |||

Db 543 LMTISFVFMIFSGLLVNLKTVVPWLSWLQYLSIPRYGYAALQHNEFLGQNFCPGLNVT 602

Qy 600 NNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655  
|| |:| |||:| |||:|||||:|

Db 603 NNTCSYAICTGEEFLTNQGIDISPWGLWKNHVALACMIVIFLTIAYLKLLFLKKFS 658

## RESULT 11

## ABCG2\_PIG

ID ABCG2\_PIG Reviewed; 656 AA.

AC Q8MIB3;

DT 21-JUN-2005, integrated into UniProtKB/Swiss-Prot.

DT 01-OCT-2002, sequence version 1.

DT 15-JAN-2008, entry version 31.

DE ATP-binding cassette sub-family G member 2 (Brain multidrug resistance protein) (CD338 antigen).

GN Name=ABCG2; Synonyms=BMDP;

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;

OC Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.

RX MEDLINE=22050127; PubMed=12054514; DOI=10.1016/S0006-291X(02)00376-5;

RA Eisenblaetter T., Galla H.-J.;

RT "A new multidrug resistance protein at the blood-brain barrier.";

RL Biochem. Biophys. Res. Commun. 293:1273-1278(2002).

CC -!- FUNCTION: Xenobiotic transporter that may play an important role  
CC in the exclusion of xenobiotics from the brain. May be involved in  
CC brain-to-blood efflux (By similarity).

CC -!- SUBUNIT: Monomer or homodimer; disulfide-linked (By similarity).

CC -!- SUBCELLULAR LOCATION: Cell membrane; Multi-pass membrane protein  
CC (By similarity).

CC -!- TISSUE SPECIFICITY: High expression in brain, kidney and lung.  
CC Also expressed in livere, colon, small intestine, heart, skeletal  
CC muscle, spleen, stomach and pancreas.

CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCG (White)  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 ABC transporter domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AJ420927; CAD12785.1; -; mRNA.  
 DR PIR; JC7860; JC7860.  
 DR RefSeq; NP\_999175.1; -.  
 DR UniGene; Ssc.64; -.  
 DR GeneID; 397073; -.  
 DR KEGG; ssc:397073; -.  
 DR GO; GO:0005886; C:plasma membrane; IEA:UniProtKB-SubCell.  
 DR InterPro; IPR003593; AAA+\_ATPase\_core.  
 DR InterPro; IPR013525; ABC\_2\_trans.  
 DR InterPro; IPR003439; ABC\_transp\_like.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 PE 2: Evidence at transcript level;  
 KW ATP-binding; Glycoprotein; Membrane; Nucleotide-binding;  
 KW Transmembrane; Transport.  
 FT CHAIN 1 656 ATP-binding cassette sub-family G member  
 FT 2.  
 FT /FTId=PRO\_0000093389.  
 FT TOPO\_DOM 1 394 Cytoplasmic (Potential).  
 FT TRANSMEM 395 415 Potential.  
 FT TOPO\_DOM 416 429 Extracellular (Potential).  
 FT TRANSMEM 430 450 Potential.  
 FT TOPO\_DOM 451 478 Cytoplasmic (Potential).  
 FT TRANSMEM 479 498 Potential.  
 FT TOPO\_DOM 499 507 Extracellular (Potential).  
 FT TRANSMEM 508 530 Potential.  
 FT TOPO\_DOM 531 536 Cytoplasmic (Potential).  
 FT TRANSMEM 537 557 Potential.  
 FT TOPO\_DOM 558 631 Extracellular (Potential).  
 FT TRANSMEM 632 652 Potential.  
 FT TOPO\_DOM 653 656 Cytoplasmic (Potential).  
 FT DOMAIN 37 286 ABC transporter.  
 FT NP\_BIND 80 87 ATP (Potential).  
 FT CARBOHYD 597 597 N-linked (GlcNAc . . .) (Potential).  
 FT CARBOHYD 601 601 N-linked (GlcNAc . . .) (Potential).  
 SQ SEQUENCE 656 AA; 72392 MW; 118ADD5B53D9D67F CRC64;

Query Match 85.0%; Score 2849.5; DB 1; Length 656;



## Q38JL0\_CANFA

ID Q38JL0\_CANFA Unreviewed; 655 AA.  
AC Q38JL0;  
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.  
DT 22-NOV-2005, sequence version 1.  
DT 08-APR-2008, entry version 31.  
DE Breast cancer resistance protein.  
GN Name=BCRP;  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae;  
OC Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Placenta;  
RA Otto A., Gabel G., Honscha K.U.;  
RT "cMXR mediated chemoresistance in canine mammary cancer.";  
RL Submitted (SEP-2005) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; DQ222459; ABB03737.1; -; mRNA.  
DR RefSeq; NP\_001041486.1; -.  
DR UniGene; Cfa.9822; -.  
DR Ensembl; ENSCAFG00000009638; Canis familiaris.  
DR GeneID; 478472; -.  
DR KEGG; cfa:478472; -.  
DR GO; GO:0016021; C:integral to membrane; IEA:UniProtKB-KW.  
DR GO; GO:0005524; F:ATP binding; IEA:InterPro.  
DR GO; GO:0016887; F:ATPase activity; IEA:InterPro.  
DR GO; GO:0006810; P:transport; IEA:UniProtKB-KW.  
DR InterPro; IPR003593; AAA+\_ATPase\_core.  
DR InterPro; IPR013525; ABC\_2\_trans.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR Pfam; PF01061; ABC2\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
PE 2: Evidence at transcript level;  
KW ATP-binding; Membrane; Nucleotide-binding; Transmembrane; Transport.  
SQ SEQUENCE 655 AA; 72718 MW; 0C2E9EDBE0A07DF3 CRC64;

Query Match 83.2%; Score 2789; DB 2; Length 655;  
Best Local Similarity 82.7%; Pred. No. 2.2e-174;  
Matches 544; Conservative 48; Mismatches 60; Indels 6; Gaps 3;

Qy	1	MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE	60
		:      :    :	
Db	1	MSSNNDPVCIPMSQRSTNDLSRMTSNDLKTSTEVAVLSFHNIIYRVKVKSGFLLGRKTVE	60
Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:      :	
Db	61	KEILTNINGVMRPGLNAILGPTGGSKSSLLDVLAARKDPHGLSGDVLINGAPRPANFKCN	120
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
		:	
Db	121	SGYVVQDDVVMGTLTVRENLQFSAALRLPTTTTSHEKNERINKVIQQGLDKVADSKVGT	180
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSSTANAVLLLLKRMSKQGRTIIF	240
		:                        :	
Db	181	QFIRGVSGGERKRTSIGMELITDPAILFLDEPTTGLDSSSTANAVLLLLKRMSEQGRTIIF	240
Qy	241	SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:                        :	
Db	241	SIHQPRYSIFKLFDSLTLLAAGKLMFHGPAQEALGFFASVGYHCEPYNNPADFFLDVING	300
Qy	301	DSTAVALNREE---DFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEK	357
		:            :             :       :            :            :	
Db	301	DSSAVVLNREDQEGEVKVTE--EPSKRGTPFIERIAEFYANSDFCRKTKEELDQLSKSQK	358
Qy	358	KKKITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLK	417
		:    :       :            :                                  :      :    :	
Db	359	RKS-SAFKEITYATSFCQQLKWISKRSFKNLLGNPQASIAQIIIVTVILGLVLGAIFYDLK	417
Qy	418	NDSTGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSD	477
		:	
Db	418	NDSTGIQNRSGVLFFLTNQCFSVSAVELLVVEKKLFIHEYISGYRVSSYFFGKLLSD	477
Qy	478	LLPMTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVA	537
		:    :          :      :                        :	
Db	478	LLPMRMLPSIIFTCIIYFLLGLKPVVEAFFIMMFTLMMVAYSASSMALAIAAGQSVVSIA	537
Qy	538	TLLMTICFVFMFMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFPCPLNA	597
		:                    :    :                :	
Db	538	TLLMTITFVFMFMIFSGLLVNLRVGPWLSWLQYLSIPRYGYAALQYNEFLGQNFPCGVNV	597
Qy	598	TGNNPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
		:            :    :    :	
Db	598	TTNNTCSYAICTGEEFLLNQGIELSPWGLWKNHVALGCMIVIFLTIAYLKLLFLKKYS	655

AC Q7TMS5; Q9R004; Q9Z1T0;  
DT 21-JUN-2005, integrated into UniProtKB/Swiss-Prot.  
DT 01-OCT-2003, sequence version 1.  
DT 08-APR-2008, entry version 43.  
DE ATP-binding cassette sub-family G member 2 (Breast cancer resistance  
DE protein 1 homolog) (CD338 antigen).  
GN Name=Abcg2; Synonyms=Abcp, Bcrp1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.  
RC STRAIN=FVB; TISSUE=Liver;  
RX MEDLINE=99413474; PubMed=10485464;  
RA Allen J.D., Brinkhuis R.F., Wijnholds J., Schinkel A.H.;  
RT "The mouse Bcrp1/Mxr/Abcp gene: amplification and overexpression in  
RT cell lines selected for resistance to topotecan, mitoxantrone, or  
RT doxorubicin.";  
RL Cancer Res. 59:4237-4241(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic stem cell;  
RX PubMed=15489334; DOI=10.1101/gr.2596504;  
RG The MGC Project Team;  
RT "The status, quality, and expansion of the NIH full-length cDNA  
RT project: the Mammalian Gene Collection (MGC).";  
RL Genome Res. 14:2121-2127(2004).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA] OF 511-657.  
RC STRAIN=C57BL/6J; TISSUE=Placenta;  
RX MEDLINE=99065313; PubMed=9850061;  
RA Allikmets R., Schriml L.M., Hutchinson A., Romano-Spica V., Dean M.;  
RT "A human placenta-specific ATP-binding cassette gene (ABCP) on  
RT chromosome 4q22 that is involved in multidrug resistance.";  
RL Cancer Res. 58:5337-5339(1998).  
RN [4]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=20493324; PubMed=11036110; DOI=10.1093/jnci/92.20.1651;  
RA Jonker J.W., Smit J.W., Brinkhuis R.F., Maliepaard M., Beijnen J.H.,  
RA Schellens J.H., Schinkel A.H.;  
RT "Role of breast cancer resistance protein in the bioavailability and  
RT fetal penetration of topotecan.";  
RL J. Natl. Cancer Inst. 92:1651-1656(2000).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=21424790; PubMed=11533706; DOI=10.1038/nm0901-1028;  
RA Zhou S., Schuetz J.D., Bunting K.D., Colapietro A.M., Sampath J.,



RA Morris J.J., Lagutina I., Grosveld G.C., Osawa M., Nakauchi H.,  
 RA Sorrentino B.P.;  
 RT "The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of  
 RT stem cells and is a molecular determinant of the side-population  
 RT phenotype.";  
 RL Nat. Med. 7:1028-1034(2001).  
 CC -!- FUNCTION: Xenobiotic transporter that may play an important role  
 CC in the exclusion of xenobiotics from the brain. May be involved in  
 CC brain-to-blood efflux (By similarity). May play a role in early  
 CC stem cell self-renewal by blocking differentiation.  
 CC -!- SUBUNIT: Monomer or homodimer; disulfide-linked (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cell membrane; Multi-pass membrane protein  
 CC (By similarity).  
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney. Lower expression  
 CC in liver, colon, heart, spleen, and placenta.  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCG (White)  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 ABC transporter domain.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF140218; AAD54216.1; -; mRNA.  
 DR EMBL; BC053730; AAH53730.1; -; mRNA.  
 DR EMBL; AF103875; AAD09189.1; -; mRNA.  
 DR RefSeq; NP\_036050.1; -.  
 DR UniGene; Mm.333096; -.  
 DR PhosphoSite; Q7TMS5; -.  
 DR Ensembl; ENSMUSG00000029802; Mus musculus.  
 DR GeneID; 26357; -.  
 DR KEGG; mmu:26357; -.  
 DR MGI; MGI:1347061; Abcg2.  
 DR ArrayExpress; Q7TMS5; -.  
 DR GermOnline; ENSMUSG00000029802; Mus musculus.  
 DR InterPro; IPR003593; AAA+\_ATPase\_core.  
 DR InterPro; IPR013525; ABC\_2\_trans.  
 DR InterPro; IPR003439; ABC\_transp\_like.  
 DR Pfam; PF01061; ABC2\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
 PE 2: Evidence at transcript level;  
 KW ATP-binding; Glycoprotein; Membrane; Nucleotide-binding;  
 KW Transmembrane; Transport.  
 FT CHAIN 1 657 ATP-binding cassette sub-family G member  
 FT 2.  
 FT /FTId=PRO\_0000093388.

Query Match 82.4%; Score 2762; DB 1; Length 657;  
Best Local Similarity 81.5%; Pred. No. 1.3e-172;  
Matches 536; Conservative 52; Mismatches 66; Indels 4; Gaps 3;

[http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7\\_142909\\_us-09-961-086a-1.rup&ItemType=4&startByte=0](http://es/ScoreAccessWeb/GetItem.action?AppId=099610...7_142909_us-09-961-086a-1.rup&ItemType=4&startByte=0) (34 of 40)9/22/2008 12:01:40 PM

Qy 360 KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIIVTVVLGLVIGAIYFGLKND 419  
: ||| | |||||:::|||||||:|:|:|:|:|:| | |  
Db 360 GTSAFKEPVYVTSFCHQLRWIARRSFKNLLGNPQASVAQLIVTVILGLIIGAIYFDLKDYD 419

Qy 420 STGIQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL 479  
: |:|||||:|||||:|||||:|||||:||||| | |:  
Db 420 AAGMQNRAGVLFFLTNNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFFGKVMSDL 479

Qy 480 PMTMLPSIIFTTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL 539  
|| |||:||||:||||| |||:||||:||||:||||| |||||  
Db 480 PMRFLPSVIFTCVLYFMLGLKKTVD AFFIMMFTLIMVAYTASSMALAIATGQSVVSVATL 539

Qy 540 LMTICFVFM MIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATG 599  
|||| ||||:||||| || |||||:|||||:||||| ||| | |  
Db 540 LMTIAFVFMMLFSGLLVNLRITGPWLSWLQYFSIPRYGFTALQYNEFLGQEFNVT 599

Qy 600 NNPC--NYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS 655  
|: | :|| ||| |||: |||:|||||:|||||:|||||  
Db 600 NSTCVNSYAICTGNEYLINQGIELSPWGLWKNHVALACMIIFLTIAYLKLLFLKKYS 657

## RESULT 14

## ABCG2\_RAT

ID ABCG2\_RAT Reviewed; 657 AA.  
AC Q80W57; Q80ST1; Q80UR3; Q80XF3;  
DT 21-JUN-2005, integrated into UniProtKB/Swiss-Prot.  
DT 01-JUN-2003, sequence version 1.  
DT 08-APR-2008, entry version 39.  
DE ATP-binding cassette sub-family G member 2 (Breast cancer resistance  
DE protein 1 homolog) (CD338 antigen).  
GN Name=Abcg2; Synonyms=Bcrp1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidea; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RX PubMed=12819005;  
RA Shimano K., Satake M., Okaya A., Kitanaka J., Kitanaka N.,  
RA Takemura M., Sakagami M., Terada N., Tsujimura T.;  
RT "Hepatic oval cells have the side population phenotype defined by  
RT expression of ATP-binding cassette transporter ABCG2/BCRP1.";  
RL Am. J. Pathol. 163:3-9(2003).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA], GLYCOSYLATION, SUBCELLULAR LOCATION, AND  
RP TISSUE SPECIFICITY.  
RC STRAIN=Wistar; TISSUE=Brain capillary;

RX PubMed=15255930; DOI=10.1111/j.1471-4159.2004.02537.x;  
RA Hori S., Ohtsuki S., Tachikawa M., Kimura N., Kondo T., Watanabe M.,  
RA Nakashima E., Terasaki T.;  
RT "Functional expression of rat ABCG2 on the luminal side of brain  
RT capillaries and its enhancement by astrocyte-derived soluble  
RT factor(s).";  
RL J. Neurochem. 90:526-536(2004).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA].  
RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
RA Yabuuchi H., Ishikawa T.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE [MRNA] OF 506-657.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain endothelium;  
RA Zhang W., Stanimirovic D.B.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Xenobiotic transporter that may play an important role  
CC in the exclusion of xenobiotics from the brain. May be involved in  
CC brain-to-blood efflux (By similarity).  
CC -!- SUBUNIT: Monomer or homodimer; disulfide-linked (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cell membrane; Multi-pass membrane protein  
CC (By similarity).  
CC -!- TISSUE SPECIFICITY: Highly expressed in brain capillary, kidney  
CC and small intestine. Lower expression in heart. Preferentially  
CC expressed (at protein level) on the luminal membrane of brain  
CC capillaries, in kidney and small intestine.  
CC -!- PTM: N-glycosylated in brain capillary, kidney and small intestine  
CC but not in heart.  
CC -!- SIMILARITY: Belongs to the ABC transporter family. ABCG (White)  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 ABC transporter domain.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AB094089; BAC75666.1; -; mRNA.  
DR EMBL; AB105817; BAC76396.1; -; mRNA.  
DR EMBL; AY089996; AAM09106.1; -; mRNA.  
DR EMBL; AY089997; AAM09107.1; -; mRNA.  
DR EMBL; AY089998; AAM09108.1; -; mRNA.  
DR EMBL; AY274118; AAP23237.1; -; mRNA.  
DR RefSeq; NP\_852046.1; -.  
DR UniGene; Rn.13131; -.  
DR Ensembl; ENSRNOG00000007041; Rattus norvegicus.  
DR GeneID; 312382; -.  
DR KEGG; rno:312382; -.  
DR RGD; 631345; Abcg2.  
DR ArrayExpress; Q80W57; -.

DR GermOnline; ENSRNOG00000007041; Rattus norvegicus.  
DR GO; GO:0005886; C:plasma membrane; IEA:UniProtKB-SubCell.  
DR InterPro; IPR003593; AAA+ATPase\_core.  
DR InterPro; IPR013525; ABC\_2\_trans.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR Pfam; PF01061; ABC2\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR ProDom; PD000006; ABC\_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
PE 1: Evidence at protein level;  
KW ATP-binding; Glycoprotein; Membrane; Nucleotide-binding;  
KW Transmembrane; Transport.  
FT CHAIN 1 657 ATP-binding cassette sub-family G member  
FT 2.  
FT /FTId=PRO\_0000093390.  
FT TOPO\_DOM 1 395 Cytoplasmic (Potential).  
FT TRANSMEM 396 416 Potential.  
FT TOPO\_DOM 417 428 Extracellular (Potential).  
FT TRANSMEM 429 449 Potential.  
FT TOPO\_DOM 450 477 Cytoplasmic (Potential).  
FT TRANSMEM 478 498 Potential.  
FT TOPO\_DOM 499 506 Extracellular (Potential).  
FT TRANSMEM 507 527 Potential.  
FT TOPO\_DOM 528 535 Cytoplasmic (Potential).  
FT TRANSMEM 536 556 Potential.  
FT TOPO\_DOM 557 632 Extracellular (Potential).  
FT TRANSMEM 633 653 Potential.  
FT TOPO\_DOM 654 657 Cytoplasmic (Potential).  
FT DOMAIN 48 285 ABC transporter.  
FT NP\_BIND 79 86 ATP (Potential).  
FT CARBOHYD 596 596 N-linked (GlcNAc. . .) (Potential).  
FT CARBOHYD 600 600 N-linked (GlcNAc. . .) (Potential).  
FT CONFLICT 363 365 AFR -> PFK (in Ref. 1; BAC75666).  
FT CONFLICT 431 431 F -> L (in Ref. 1; BAC75666).  
FT CONFLICT 492 492 I -> L (in Ref. 3; AAM09106/AAM09107/  
FT AAM09108).  
FT CONFLICT 502 502 T -> L (in Ref. 1; BAC75666).  
FT CONFLICT 510 510 M -> R (in Ref. 1; BAC75666).  
SQ SEQUENCE 657 AA; 72961 MW; C975C61A08489027 CRC64;

Query Match 82.2%; Score 2754; DB 1; Length 657;  
Best Local Similarity 81.0%; Pred. No. 4.4e-172;  
Matches 533; Conservative 52; Mismatches 69; Indels 4; Gaps 3;

Qy 1 MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLGSGFLPCRKPVE 60  
| | | | | : | : | | | | | : | : | | | | | : | | | | | | | | |  
Db 1 MSSSNDHVLVPMSSQRNKNGLPGMSSRGARTLAEGDVLSFHHITYRVKVKSGFL-VRKTAE 59

Qy	61	KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN	120
		:	
Db	60	KEILSDINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPRGLSGDVLINGAPQPANFKCS	119
Qy	121	SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT	180
Db	120	SGYVVQDDVVMGTLTVRENLQFSAALRLPKAMKTHEKNERINTIikelGLDKVADSKVGT	179
Qy	181	QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	240
		:	
Db	180	QFTRGISGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF	239
Qy	241	SIHQPRYSIFKLFDSLTLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING	300
		:     :                 :	
Db	240	SIHQPRYSIFKLFDSLTLASGKLMFHGPAQKALEYFASAGYHCEPYNNPADFFLDVING	299
Qy	301	DSTAVALNR-EEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKK	359
		:        :   :  :     : :       : :           :	
Db	300	DSSAVMLNRGEQDHEANKTEEPSKREKPIIENLAEFYINSTIYGETKAELDQLPVAQKKK	359
Qy	360	KITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKND	419
		:  :         : :       : :      : :	
Db	360	GSSAFREPVYVTSFCHQLRWIARRSFKNLLGNPQASVAQLIVTVILGLIIGALYFGLKND	419
Qy	420	STGIQNRAGVLFFLTNQCFSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLL	479
		:           :	
Db	420	PTGMQNRAGVFFFLTNTQCFSTSVSAVELFVVEKKLFIHEYISGYRVSSYFFGKLVSDLL	479
Qy	480	PMTMLPSIIFTTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATL	539
		: : :        :  : : : : :	
Db	480	PMRFLPSVIYTCILYFMLGLKRTVEAFFIMMFTLIMVAYTASSMALAIAAGQSVVSVATL	539
Qy	540	LMTICFVFMMLFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFPCPLNATG	599
		:	
Db	540	LMTISFVFMMLFSGLLVNLRTIGPWLSWLQYFSIPRYGFTALQHNEFLGQEFPCPLNVTM	599
Qy	600	NNPC--NYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS	655
		:   :       : :       : :       : :	
Db	600	NSTCVNSYTICTGNDYLINQGIDLSPWGLWRNHVALACMIIFLTIAYLKLLFLKKYS	657

RESULT 15  
Q28BS4\_XENTR  
ID Q28BS4\_XENTR Unreviewed; 661 AA.  
AC Q28BS4;  
DT 04-APR-2006, integrated into UniProtKB/TrEMBL.  
DT 04-APR-2006, sequence version 1.  
DT 08-APR-2008, entry version 19.

Query Match 69.9%; Score 2343; DB 2; Length 661;  
Best Local Similarity 69.2%; Pred. No. 4.8e-145;  
Matches 456; Conservative 81; Mismatches 102; Indels 20; Gaps 5;

Qy	6	VEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVEKEILS	65
		::             :     : :    :        :   :	
Db	10	VQILDPTVNGEVK-----KKGRKKTLSGAVLSFYNNINYKVVKVKSGLICCRKVTERVILN	63
Qy	66	NINGIMKPGLNAILGPTGGGKSSLLDVLAAARKDPSSLSDVLLINGAPRPANFKCNSGYVV	125

Db	64	DVNGIMKPGLNAILGPTGSGKSSLLDVLAARKDPNGLSGQVLVDGEPQPSNFKCLSGYVV	123
Qy	126	QDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGTQFIRG	185
Db	124	QDDVVMGTLSIRENLQFSAALRLPRSVKQKEKDERINQVIKELGLTKVADSKVGTQFIRG	183
Qy	186	VSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIFS IHQP	245
Db	184	VSGGERKRTNIGMELITDPGILFLDEPTTGLDASTANAVLLLLKRMSRQGKTIIFS IHQP	243
Qy	246	RYSIFKLFDSLTL LASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIINGDSTAV	305
Db	244	RYSIFRLFDSLTL LAGGRLLFHGPSRDALDYFTGLGYECESHNNPADFFLDIINGDSTAV	303
Qy	306	ALNREEDFKATEIIEPSKQ-----DKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK	360
Db	304	ALNKLED---VELENEQKEVNDNGSKTVVENLSEQFCTTSYYLETKAELEKMSLGKKIKS	360
Qy	361	ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS	420
Db	361	NFFARQITYNTSFLHQLKWVCKRSFKNLWRNPQASIAQVMVTLVLALIVGAIFFGVKEDV	420
Qy	421	TGIQNRAGVLFFLT TNQCFSSVSAVELFVVEKKLFIHEYISGYRVSSYFLGKLLSDLLP	480
Db	421	SGIQNRVGS LFFVT TNQCFSSVSAIELFIVEKKIFIHEYISGYRLSAYFFAKLFTDLLP	480
Qy	481	MTMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL	540
Db	481	MRTLPSIIFTSVIYFMIGFKATAGAFFTMMFTLMMIAYTAASMALAVAAGQDVVAVANLL	540
Qy	541	MTICFVFMMIFSGLLVNLT TIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPLNAT--	598
Db	541	MTICFVFMIIFSGLLVNLT SVMDWISWLKYFSIPRYGLTALQINEFTNLNFCNGLNTTIQ	600
Qy	599	GNNPCN----YATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLT IAYLKLLFLKK	653
Db	601	GNPNCTGSSPFGTCTGEEYLT VQGIDFSTWGLWQNLHALACMIAIFLT IAYLKLYFMKK	659

Search completed: September 18, 2008, 22:07:02  
Job time : 414 secs